

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,300

Open access books available

130,000

International authors and editors

155M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Selected Issues and Constraints of Image Matching in Terrain-Aided Navigation: A Comparative Study

Piotr Turek, Stanisław Grzywiński and Witold Buzantowicz

Abstract

The sensitivity of global navigation satellite systems to disruptions precludes their use in conditions of armed conflict with an opponent possessing comparable technical capabilities. In military unmanned aerial vehicles (UAVs) the aim is to obtain navigational data to determine the location and correction of flight routes by means of other types of navigational systems. To correct the position of an UAV relative to a given trajectory, the systems that associate reference terrain maps with image information can be used. Over the last dozen or so years, new, effective algorithms for matching digital images have been developed. The results of their performance effectiveness are based on images that are fragments taken from source files, and therefore their qualitatively identical counterparts exist in the reference images. However, the differences between the reference image stored in the memory of navigation system and the image recorded by the sensor can be significant. In this paper modern methods of image registration and matching to UAV position refinement are compared, and adaptation of available methods to the operating conditions of the UAV navigation system is discussed.

Keywords: digital image processing, image matching, terrain-aided navigation, unmanned aerial vehicle, cruise missile

1. Introduction

Global navigation satellite systems are widely used in both civil and military technology areas. The advantage of such systems is very high accuracy in determining the coordinates, however, the possibility of easy interference precludes their use in conditions of armed conflict with an opponent equipped with comparable technical capabilities. In the case of military autonomous unmanned aerial vehicles (UAVs), in particular cruise missiles (CM), the aim is therefore to determine navigation data for specifying the position and correcting the flight paths by means of other types of navigation and self-guidance systems.

Such systems are usually based on inertial navigation systems (INS) which use accelerometers, angular rate gyroscopes and magnetometers to provide relatively accurate tracking of an object's position and orientation in space. However, they are exposed to drift and systematic errors of sensors, hence the divergence between the actual and the measured position of the object is constantly increasing with time. This results in a significant navigational error.

Therefore, two types of systems designed to correct the position of an object in relation to a given trajectory are normally used in the solutions of the UAV/CM navigation and self-guidance systems. The first group contains systems whose task is to determine the position on the basis of data obtained from radio altimeters, related to reference height maps. Such systems include, for example: TERCOM (terrain contour matching), used in Tomahawk cruise missiles, SITAN (Sandia inertial terrain-aided navigation), using terrain gradients as input for the modified extended Kalman filter (EKF) estimating the position of the object, and VATAN (Viterbi-algorithm terrain-aided navigation), a version of the system based on the Viterbi algorithm and characterised – in relation to the SITAN system – with lower mean square error of position estimation [1–5]. The main disadvantage of these solutions is the active operation of measuring devices, which reveals the position of the object in space and eliminates the advantages associated with the use of a passive (and therefore undetectable) inertial navigation system. The second group consists of systems associating reference terrain maps with image information obtained by means of visible, residual or infrared light cameras [6, 7]. Such systems include the American DSMAC (digital scene matching area correlator), also used in Tomahawk missiles [8, 9], and its Russian counterpart used in Kalibr (aka Club) missiles. Their advantage is both the accuracy of positioning and the secrecy (understood as passivity) of operation.

Due to the dynamic development of UAVs/CMs equipped with navigation systems operating independently of satellite systems and a number of problems associated with the implementation of the discussed issue, the assessment on the sensitivity of the selected methods to environmental conditions and constraints in the measurement systems, which often negatively affect the results obtained, has been carried out. The essence of the work is to consider issues related to the processing of image information obtained from optical sensors carried by UAV/CM and its association with terrain reference images. In particular, issues of the correctness of image data matching and the limitations of the possibilities of their similarities' assessment are considered. The article compares modern image matching methods assuming real conditions for obtaining information. The main goal set by the authors is to verify selected algorithms, identify the key aspects determining the effectiveness of their operation and indicate potential directions of their development.

2. State of the art

The operation of classic object identification algorithms, indicating the similarities between the recorded and reference images (the so-called *patterns*), is mainly based on the use of correlation methods. These algorithms, although effectively implemented in solutions to typical technical problems, are insufficiently effective in the case of topographic navigation. It is related to, inter alia, the limitations and conditions in the measurement system, environmental conditions and characteristics of the detected objects, which have a strong negative impact on the obtained correlation results. This disqualifies the possibility of their direct use in the tasks of matching reference terrain maps with the acquired image information.

A particularly significant obstacle is the fact that the sensory elements of navigation systems installed on UAV/CM record image data in various environmental and lighting conditions [10]. Frequently, reference data of high informative value, due to various conditions, constitute a pattern of little use or even lead to incorrect results. This is the case, for example, when the reconnaissance is



Figure 1.
Images of the same fragment of the Earth's surface taken under different weather and lighting conditions.

conducted in different weather conditions than those in which the UAV/CM mission takes place (**Figure 1**). Therefore, image feature matching becomes a complex issue. The conditions related to the image recording parameters, e.g. variable view angle, maintaining scale or using various types of sensors, turn out to be equally important.

Image matching methods began to be strongly developed with the dissemination of digital image in technology. Initially, the classical Fourier and correlation methods were used. However, these methods did not allow for successful multi-modal, multi-temporal, and multi-perspective matching of different images. The taxonomy of the classical methods used in the image matching process was presented in the early 1990s [11]. The image feature space, considered as a source of information necessary for image matching, was defined and local variations in the image were identified as the greatest difficulty in the matching process. In the 21st century, further development of methods based on the features of the image continued [12]. It should be emphasised that most image matching methods based on image features include four consecutive basic steps: feature detection, feature matching, model estimation of mutual image transformation and final transformation of the analysed image. These methods became an alternative to the correlation and Fourier methods. For over a dozen years, new, effective algorithms for processing and matching digital images have been developed, using statistical methods based on matching local features in images [11–14], cf. **Figure 2**. Their authors point to the greater invariance of the proposed algorithms to perspective distortions, rotation, translation, scaling and lighting changes. Given their high reliability under static conditions, as well as their low sensitivity to changes in the optical system's position, including translation, orientation and scale, it is justified to conduct studies in order to verify their usefulness and effectiveness. The paper focuses on modern image matching algorithms, which can potentially be used in topographic navigation issues. It should be stressed that the problem is completely different in the indicated context. This is due to the fact that although the matched images represent the same area of the terrain, the manner and time of the recording differ significantly from each other. This is not a typical application of these algorithms, hence a limited effectiveness of their operation can be expected.

The common feature of all methods is the use of the so-called *scale space* described in [15], allowing the decimation of image data and examination of

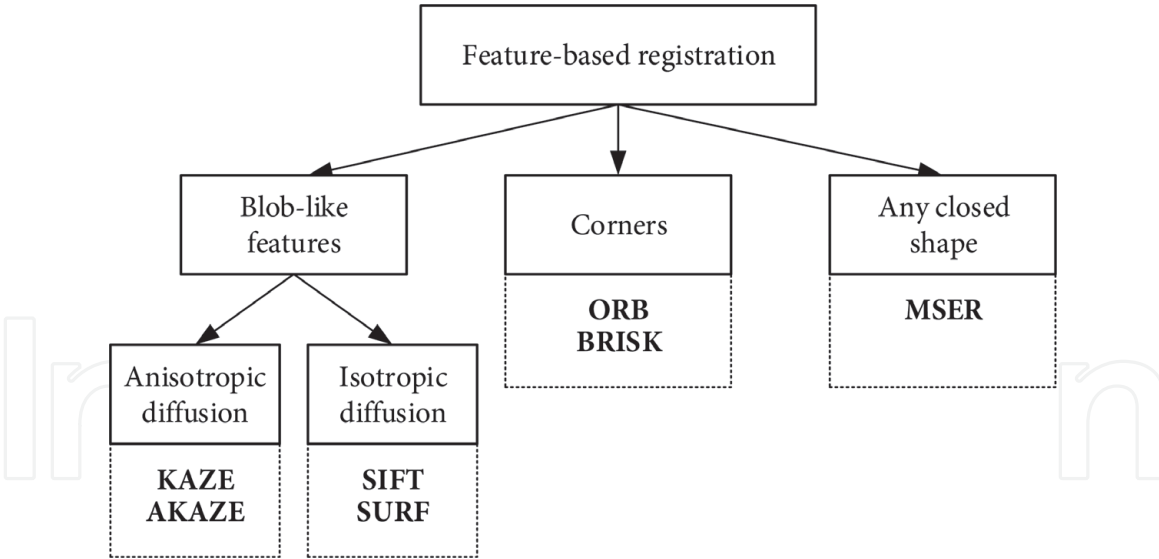


Figure 2.
Classification of the selected methods of image feature matching.

similarities between images of different scales. A significant step in the development of image matching methods based on local features was the development of the Scale-Invariant Feature Transform (SIFT) algorithm [16]. In this algorithm, the characteristics are selected locally and their position does not change while the image is scaled. Their indication is done by determining the local extremes of function $D(\hat{x})$, that is the difference between the results of image $I(x, y)$ convolution with Gaussian functions $G(x, y, \sigma)$ with different values of scale parameter σ :

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (1)$$

where

$$\begin{aligned} D = D(x, y, \sigma) &= [G_{\sigma_1}(x, y) - G_{\sigma_2}(x, y)] \star I(x, y) = \\ &= \frac{1}{\sqrt{2\pi}} \left[\frac{1}{\sigma_1} e\left(-\frac{x^2 y^2}{2\sigma^2}\right) - \frac{1}{\sigma_2} e\left(-\frac{x^2 y^2}{2\sigma^2}\right) \right] \star I(x, y) \end{aligned} \quad (2)$$

A more numerically efficient version of the SIFT algorithm, called Speeded-Up Robust Features (SURF) is based on the so-called *integral images* [17]. Both methods use the basic processing steps described in [12]. Additionally, in order to ensure the effectiveness of feature detection in images of different resolutions, a scale space, consisting of octaves which represent the series of responses of a convolutional filter with a variable size, was introduced.

Simply put, the detection of the characteristic point is based on the use of the determinant of a Hessian matrix $\det(H)$. In the case of SURF, the second-order derivatives of the Gaussian function G approximated by the box filters B_{xx} , B_{xy} , B_{yy} , and the integral image are also used [18]. The Hessian matrix in these methods takes the form

$$H = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (3)$$

where

$$\begin{aligned} L_{xx}(x, y, \sigma) &= \frac{\partial^2}{\partial x^2} G(x, y, \sigma) \star I(x, y) \approx B_{xx} \star I(x, y) \\ L_{xy}(x, y, \sigma) &= \frac{\partial^2}{\partial x \partial y} G(x, y, \sigma) \star I(x, y) \approx B_{xy} \star I(x, y) \\ L_{yy}(x, y, \sigma) &= \frac{\partial^2}{\partial y^2} G(x, y, \sigma) \star I(x, y) \approx B_{yy} \star I(x, y) \end{aligned} \quad (4)$$

The determinant of the Hessian matrix after approximation using box filters and the Frobenius norm is given as

$$\det(H) \approx B_{xx}B_{yy} - \left(\frac{9}{10}B_{xy}\right)^2 \quad (5)$$

After detecting the local extremes of $\det(H)$, similarly to $D(\hat{x})$, the location of characteristic points representing local features, called *blob-like* for the SIFT and SURF methods, is determined. In this step, the SIFT method also rejects features whose contrast is lower than the assumed threshold t by comparing $|D(\hat{x})| < t$ as well as points lying on isolated edges. This is done by comparing the value of the Hessian matrix H trace quotient, and its determinant with the curvature coefficient r :

$$\frac{\text{tr}(H)}{\det(H)} < \frac{(r+1)^2}{r} \quad (6)$$

In 2011, an alternative method to SIFT and SURF, Oriented FAST and Rotated BRIEF (ORB), was proposed [19]. The method was based on the modified Features from Accelerated Segment Test (FAST) detector [20, 21], enabling corner and edge detection, and a modified Binary Robust Independent Elementary Features (BRIEF) descriptor [22]. This approach involves changing the scale of the image on the basis of blurring with an increasing value of Gaussian filter. Despite the noise reduction and enhancing the uniformity of areas interpreted by human beings as unique (e.g. surface of the lake, wall of a building, shape of a vehicle, etc.), it causes blurring of their edges. This often leads to the inability to indicate the boundaries between areas and to define characteristic points in their neighbourhood.

The solution to this problem was proposed in the KAZE method (Japanese for “wind”) [23]. Unlike the SIFT and SURF methods, which use the Gaussian function causing isotropic diffusion of luminance to generalise the image, in the KAZE method the generalisation is based on nonlinear diffusion in consecutive octaves of the scale [24]. The anisotropic image blurring in this method depends on the local luminance distribution. Nonlinear diffusion can be presented in the following equation:

$$\frac{\partial I}{\partial t} = \text{div}[c(x, y, t)\nabla I] \quad (7)$$

The blur intensity can be adapted by the introduced conductivity function c , which is usually related to time. However, using the approach proposed in [15], parameter t is related to the image scale. Various forms of the conductivity function c were proposed in related works developing the use of nonlinear diffusion in the context of image filtration [24–26]. One of the functions used for nonlinear diffusion can be:

$$c = \exp \left(-\frac{|\nabla L_\sigma|^2}{k^2} \right) \quad (8)$$

where ∇L_σ is the gradient of the Gaussian blurred function of the original image I on the scale σ , and k is the contrast ratio.

This function allows for blurring the image while maintaining the edges of structures. As a result, more features can be detected at different image scales. However, it involves the use of a gradient, which in the case of intense image disturbance, e.g. in the form of a shadow, may cause an unfavourable (due to the subsequent detection of features) distribution of diffusion in the image.

An important stage of the considered methods is the description of a characteristic point by means of a vector containing information about its surroundings. The SIFT method uses a luminance gradient and in the SURF method the image response to horizontally and vertically oriented Haar wavelet is applied. In general, around the characteristic point in the area with a defined radius dependent on the σ scale, a certain number of cells are created and dominant values of the gradient or the responses to Haar wavelets are determined. These are the basis for calculating the so-called *feature metrics*. Finally, the dominant orientation is established. Characteristic features in the SIFT method are determined by

$$\begin{aligned} m(x, y) &= \sqrt{[L_\sigma(x+1, y) - L_\sigma(x-1, y)]^2 + [L_\sigma(x, y+1) - L_\sigma(x, y-1)]^2} \\ \theta(x, y) &= \tan^{-1} \left[\frac{L_\sigma(x, y+1) - L_\sigma(x, y-1)}{L_\sigma(x+1, y) - L_\sigma(x-1, y)} \right] \end{aligned} \quad (9)$$

where $m(x, y)$ is the gradient size, $\theta(x, y)$ is the orientation, and L_σ is the blurred image discussed above.

The SIFT method creates thereunder a gradient histogram that sums up the determined values in four cells. In analogous cells, according to the SURF method, the responses to Haar wavelets distributed along the radii in the neighbourhood of the point with an interval of $\pi/3$ are summed up. In each SURF subregion, a vector \mathbf{v} is determined:

$$\mathbf{v} = \left[\sum d_x \quad \sum d_y \quad \sum |d_x| \quad \sum |d_y| \right] \quad (10)$$

where d_x and d_y are the characteristic point's neighbourhood responses to the horizontally and vertically oriented Haar wavelets, respectively.

In the KAZE method the procedure is similar as for the SURF method with the difference that the first order derivatives from the image function are used. The point description operation is performed for all levels in the adopted scale space, thereby creating a pyramid of vectors assigned to subsequent levels containing an increasingly generalised image.

The Maximally Stable Extremal Regions (MSER) method introduced in [27] has a different approach to the detection and description of local features. In this method, regions (shapes), referred to as *maximally stable*, are selected as the characteristics of the image. The image in this method is treated as a function I which transforms

$$I : \mathcal{D} \subset \mathbb{Z}^2 \rightarrow \mathbf{S} \quad (11)$$

where \mathcal{D} is the domain of I and \mathbf{S} is its set of values, usually $\mathbf{S} = \{0, 1, \dots, 255\}$.

Regions (areas, shapes) with a specific (typically average) luminance level can be determined in the image. Region Q is understood as a subset of pixels in the

image which is a continuous subset of \mathcal{D} such that $\forall p, q \in Q$ exist sequences p, a_1, a_2, \dots, q and $pAa_1, \dots, a_iAa_{i+1}, \dots, a_nAq$, where $A \subset \mathcal{D} \times \mathcal{D}$ is the neighbourhood relation and the formula a_iAa_{i+1} is the neighbourhood between pixels a_i and a_{i+1} . The extremal region is $Q \subset \mathcal{D}$, such that $\forall p \in Q, \forall q \in \partial Q : I(p) > I(q)$ (maximum intensity region) or $I(p) < I(q)$ (minimum intensity region). The desired maximally stable extremal region (MSER) is the region $R = Q_{i^*}$, which for the sequence of extremal regions Q_1, \dots, Q_{i-1}, Q_i , nested i.e. $Q_i \subset Q_{i+1}$ and for $q(i) = |Q_{i+\Delta} \setminus Q_{i-\Delta}| / |Q_i|$ has a local minimum at i^* . Whereas $\Delta \in \mathbb{S}$ is the stability parameter and the luminance threshold. The procedure of determining MSER regions is repeated throughout the assumed σ scale space.

In the feature description stage, a vector using image moments is determined for each region. Based on the moments $m_{00}, m_{01}, \dots, m_{20}$, the centre of gravity of each MSER region and the ellipse approximating the region are determined according to the procedure described in [28]. The ellipse equation is given as

$$\frac{\left[x - x_g + \theta(y - y_g)\right]^2}{a_1(1 + \theta^2)} + \frac{\left[y - y_g + \theta(x - x_g)\right]^2}{a_2(1 + \theta^2)} - 1 = 0 \quad (12)$$

The orientation θ and the size of the ellipse defined by its a_1 and a_2 axes allow to describe the features of the region taken for comparison in the matching step. The moment m of the order $(p + q)$ of the MSER region used to determine the centre of gravity of the $C = I(x_g, y_g)$ region can be represented as follows:

$$m_{pq} = \sum_{\{x,y\} \in \mathbb{R}} x^p y^q \quad (13)$$

The use of moments and the centre of gravity is also a feature of ORB method which uses machine learning approach for corner detection. After their detection, based on the image moments, the centre of gravity C is determined for each corner according to the formula

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (14)$$

where

$$m_{pq} = \sum x^p y^q I(x, y) \quad (15)$$

On the basis of the corner's position and centre of gravity, the orientation of the feature is determined as shown in the equation:

$$\theta(x, y) = \text{atan2}(m_{01}, m_{10}) \quad (16)$$

The feature description step uses the assigned orientation to complete the binary BRIEF descriptor [22], with the condition of verifying the belonging of the point $L_\sigma(x, y)$ to the matrix \mathbf{W}_θ . It is based on the simple comparison of the pixel luminance values in the neighbourhood of the feature:

$$\tau(L_\sigma; x, y) = \begin{cases} 1 & \Leftrightarrow L_\sigma(x, y) < L_\sigma(x_1, y_1) \\ 0 & \Leftrightarrow L_\sigma(x, y) \geq L_\sigma(x_1, y_1) \end{cases} \quad (17)$$

The matrix \mathbf{W}_θ is the product of the original matrix \mathbf{W} , containing the locations of the points, which are subject to the tests, and the rotation matrix based on the determined angles $\theta(x, y)$. In such case the ORB vector describing the feature takes the form:

$$\mathbf{v}_n = \sum_{1 \leq i \leq n} 2^{i-1} \tau(L_\sigma; x_i, y_i) | (x_i, y_i) \in \mathbf{W}_\theta \quad (18)$$

The common element for the described methods is the stage of comparing the distinguished features detected on the reference and registered images. It is of fundamental importance in the field of absolute terrain position designation, because the location of the matched features is the source of determining the matrix of mutual image transformation. In this comparison, vectors describing the features in a given method, e.g. feature metric and its orientation, are taken into account.

The determination of the similarity between the feature description vectors \mathbf{v}_a and \mathbf{v}_b is based on various measures. The most commonly used are the distances defined as follows:

$$d_1(\mathbf{v}_a, \mathbf{v}_b) = \sum |\mathbf{v}_a - \mathbf{v}_b| \text{ and } d_2(\mathbf{v}_a, \mathbf{v}_b) = \sum (\mathbf{v}_a - \mathbf{v}_b)^2 \quad (19)$$

The third frequently used norm for binary vectors is the Hamming distance given as:

$$d_3(\mathbf{v}_a, \mathbf{v}_b) = \sum \text{XOR}(\mathbf{v}_a, \mathbf{v}_b) \quad (20)$$

Another approach for matching two features is the nearest neighbour algorithm based on the ratio of the distances d_1 and d_2 . However, it should be remembered that the matching result for the described distances may vary, hence the importance of the features' detection and description step.

The final step in all the discussed methods is the statistical verification of a set of matched local features. It happens that, as a result of the initial comparison of the vectors which describe the features, mismatches resulting from the acquisition conditions described above are indicated. Therefore, after the pre-processing step, additional criteria are applied to distinguish matches from mismatches, e.g. based on the Random Sample Consensus (RANSAC) method [29]. This method allows for the estimation of a mathematical model describing the location of local features in the image provided that most of the matched points fit into this model (with the assumed maximum error). Then those points that do not fit into the estimated model are discarded in the step of determining the image transformation matrix.

3. Problem formulation

The following set is considered:

$$\mathbf{I} = \{I_i, i \in \mathbb{N}\} \quad (21)$$

Elements of \mathbf{I} are two-dimensional discrete signals (digital images) and describe the same part of the Earth's surface, but recorded at different times and therefore under different environmental and lighting conditions. Image I_j chosen from the set \mathbf{I} is treated as a reference signal, i.e. characterised by excellent structural similarity to itself. In order to compare any I_k image selected from the set \mathbf{I} with reference

image I_j , the following similarity measures were used: mean square error J_{MSE} and the related J_{PSNR} (peak-signal-to-noise ratio) and J_{SSIM} (structural similarity index measure).

The mean square error is determined by the formula:

$$J_{\text{MSE}}(I_j, I_k) = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [I_j(x, y) - I_k(x, y)]^2 \quad (22)$$

in which M and N are the width and height of images in pixels. Index J_{PSNR} is defined as:

$$J_{\text{PSNR}}(I_j, I_k) = 10 \log_{10} \left\{ \frac{\ell^2}{J_{\text{MSE}}(I_j, I_k)} \right\} \quad (23)$$

where ℓ is the range of changes of the luminance value, while the index J_{SSIM} can be described as:

$$J_{\text{SSIM}}(I_j, I_k) = \frac{(2\mu_{I_j}\mu_{I_k} + \xi_1)(2\sigma_{I_j I_k} + \xi_2)}{(\mu_{I_j}^2 + \mu_{I_k}^2 + \xi_1)(\sigma_{I_j}^2 + \sigma_{I_k}^2 + \xi_2)} \quad (24)$$

in which μ_{I_j} and μ_{I_k} are the mean values of the I_j and I_k image luminance, σ_{I_j} and σ_{I_k} are the variances of I_j and I_k , $\sigma_{I_j I_k}$ is covariance of a pair (I_j, I_k) , and $\xi_1 = (0.01\ell)^2$ and $\xi_2 = (0.03\ell)^2$ are positive constants avoiding instability when the denominator of the Eq. (24) is very close to zero [30–32].

For such defined initial conditions, the best match of the subsequent elements of the set \mathbf{I} in relation to the reference element I_j is sought, assuming that the similarity index measures of the examined pairs (I_j, I_k) are strongly undesirable, i.e.

$$J_{\text{PSNR}}(I_j, I_k) \rightarrow 0 \quad \text{and} \quad J_{\text{SSIM}}(I_j, I_k) \ll 1 \quad (25)$$

The term *best match* is understood as defining certain vectors \mathbf{v}_j and \mathbf{v}_k of values, which characterise the considered signals I_j and I_k , and then linking them in a way that makes it possible to explicitly state that the selected pair (I_j, I_k) describes the same fragment of the Earth's surface.

4. Performance analysis

In order to verify the sensitivity of the selected methods to limitations in the measurement system and environmental changes, a number of studies taking into account the actual conditions of obtaining information were conducted. Due to their difficult nature, they were performed with the use of computer simulation methods. The research was carried out in three stages. In the first stage, a detailed analysis of the test sets, using the values of the similarity indexes of the elements defined in the article, was completed. On the basis of the performed tests, special cases were selected and subjected to detailed analysis. In the further part of the study, the methods and verification of the correctness of image data matching in the scope of mutual matching of the sets presented to the analysis were compared. Finally, the influence of changes in the contrast of the acquired image on the number of features detected and the subsequent matching results was examined.

4.1 Analysis of test set elements

For the initial numerical tests, the test set **I** consisting of four elements was adopted, whereby I_0 is treated as reference. Each element of the set **I** is a 24-bit digital image with a size of 1080×1080 pixels, representing the same fragment of the Earth’s surface, with a centrally located characteristic terrain object (**Figure 3**).

The object is located in a natural environment characteristic for tundra and therefore distinguished by a rocky ground with a very low plant cover, dominated by mosses and lichens. Image I_0 (reference) was taken in the autumn and mostly brown colours, associated with the tundra soils and rock formations in this area, prevail in it. The image I_1 shows the environment in spring–summer conditions, i.e. during the growing season. Images I_2 and I_3 were taken in winter, with snow cover, whereby in the case of I_3 , there is also a strong cloud cover. Test set **I** elements’ similarity index measures, determined on the basis of the Eqs. (22)–(24) with the reference image I_0 , are presented in **Table 1** (columns 2–5).

Based on the obtained results, it can be shown that the elements constituting the test set **I** were selected so that only one of them (I_1) has a relatively high degree of

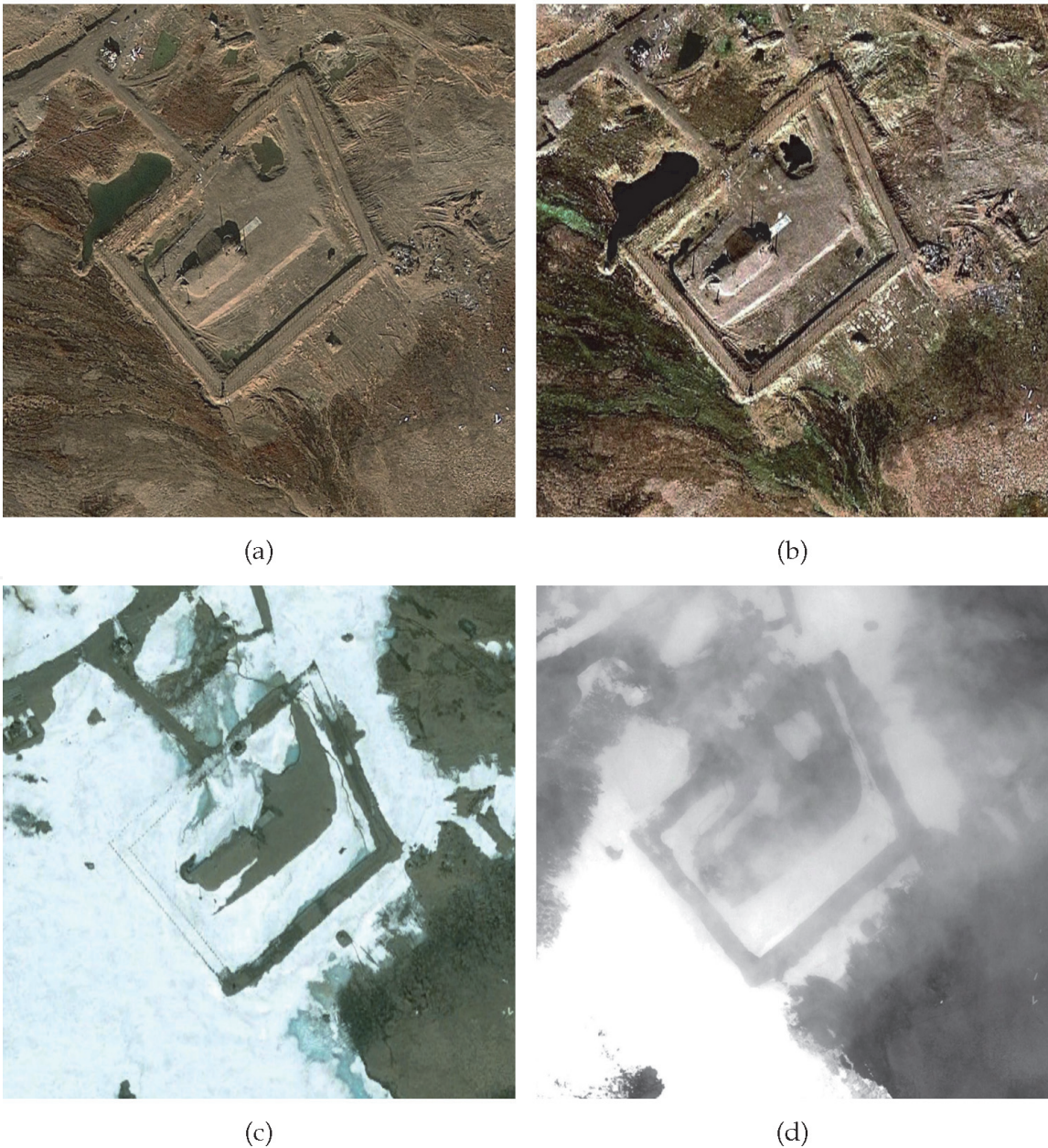


Figure 3.
Test image set **I**: (a) reference image I_0 , (b)–(d) test images I_1, I_2, I_3 (source: Google Earth).

	(I_0, I_0)	(I_0, I_1)	(I_0, I_2)	(I_0, I_3)	(I_2, I_2)	(I_2, I_3)
J_{MSE}	0	2.23E03	1.46E04	1.19E04	0	2.99E03
J_{PSNR}	∞	14.65	6.47	7.36	∞	13.37
J_{SSIM}	1	0.4373	0.0742	0.0755	1	0.3378

Table 1.
Similarity index measures for selected pairs of the set **I**.

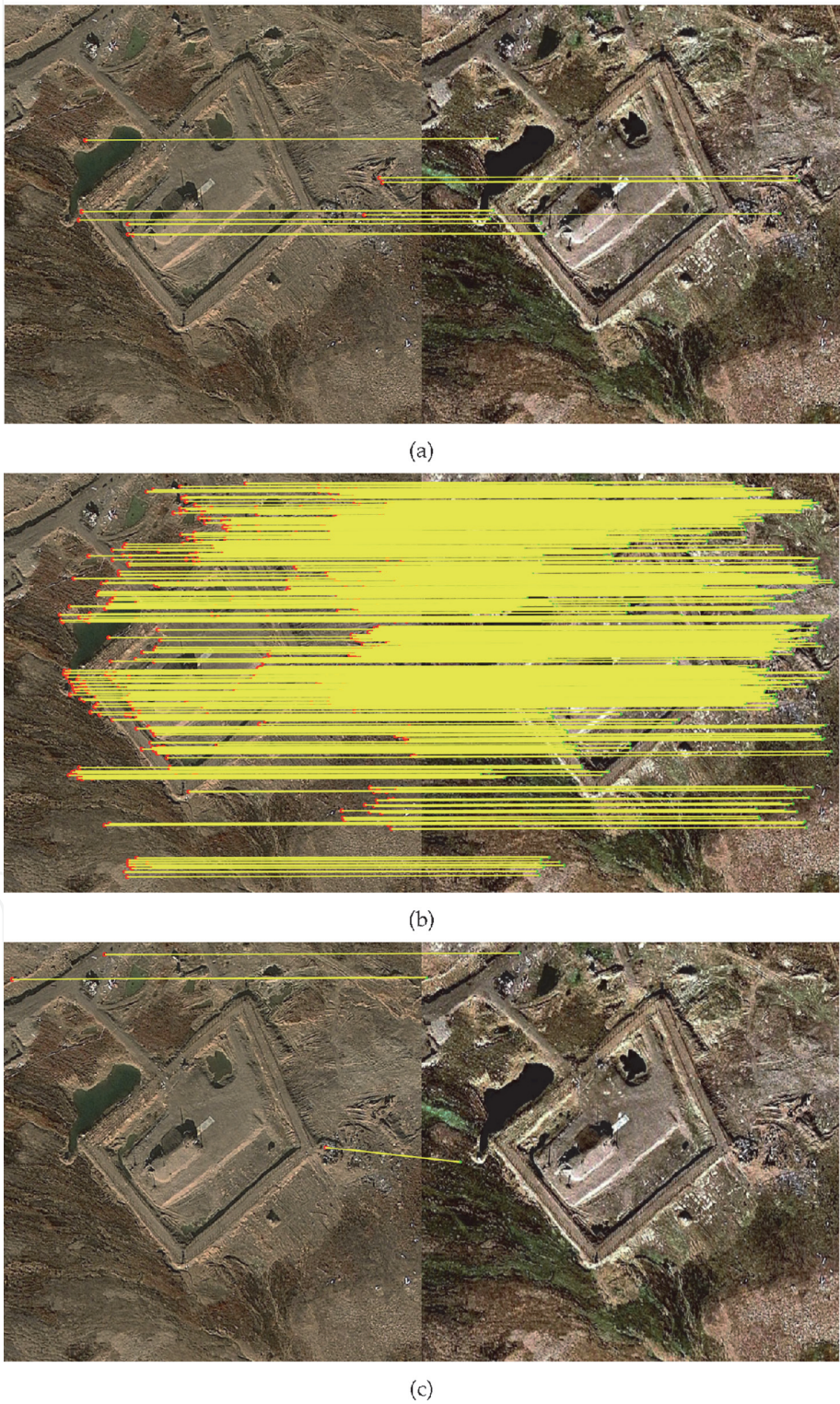


Figure 4.
Image pair (I_0, I_1) matching result for: (a) SURF, (b) KAZE, and (c) MSER method.

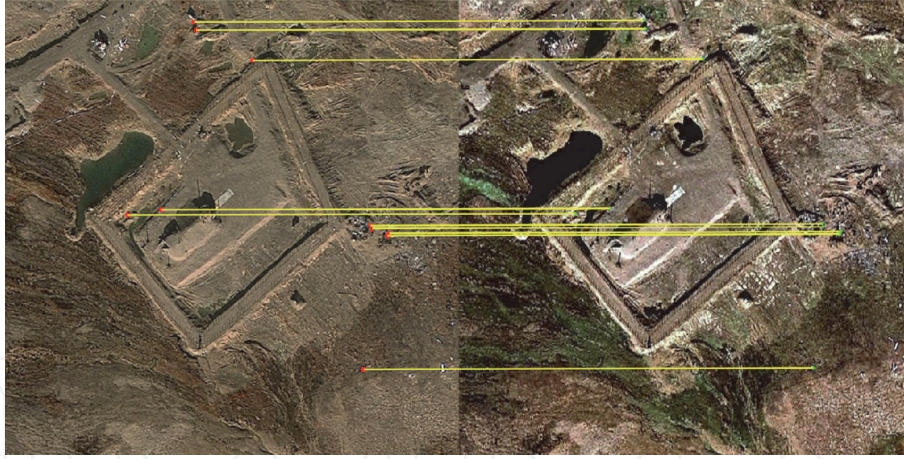


Figure 5.
Image pair (I_0, I_1) matching result for ORB method.

similarity to the reference image. The remaining items (I_2 and I_3) have unfavourable similarity index measures J_{PSNR} and J_{SSIM} , which enables the assumptions of the Eq. (25) to be met. Although in a subjective manner the image I_2 is more similar to the reference image I_0 , I_3 is characterised by more favourable J_{PSNR} and J_{SSIM} index measures.

It should be noted that I_2 and I_3 were also carefully selected. Both images show a similar arrangement of snow cover, which is reflected in the determined values of the pair similarity indexes (I_2, I_3) , cf. **Table 1**, columns 6 and 7. This is justified in practice: it may happen that the data from the reconnaissance are accurate (e.g. they take into account the snow in the area concerned, the lack of leaves on the trees in late autumn or the high water level in spring), but strong fogging or rainfall make the image I_3 obtained from UAV/CM recording systems several hours later significantly different from the reference image (in this case I_2).

4.2 Comparison of the selected methods of image feature matching

The test set **I** presented in SubSection 4.1 was examined in order to compare the effectiveness of image matching performed by algorithms using local features. The SURF, MSER, ORB and KAZE methods were taken into account. Image I_0 is a pattern, and I_1, I_2, I_3 are matched images. In the algorithms, the values of parameters proposed by their authors were used with the exception of the features' similarity threshold, which was lowered to the level of 50% due to large differences between individual elements of the set **I**. The best matching of individual features in the compared images was assumed, using the similarity measures proposed for these methods. The RANSAC method was used for the final correction of the matched features, for which an affine transformation model between images was adopted. In order to verify the effectiveness of the considered methods and the correctness of the parameters adopted in the last study, the pattern was replaced. It was assumed that I_2 is a reference image and I_3 is a matched image. The matching results of the individual test pairs of **I** are shown in **Figures 4-8** and in **Tables 2-5**.

Analysis of the matching results has shown that the selected algorithms are not effective when the matched images, despite the same content, differ significantly, cf. pair (I_0, I_3) . All methods were most effective in matching the pair (I_0, I_1) . While the SURF and MSER methods indicated mismatches for matching pairs (I_0, I_2) and (I_0, I_3) , the ORB method did not (cf. **Table 3** and **Table 4**). The KAZE method identified correctly the fragment of the image on which the corresponding features

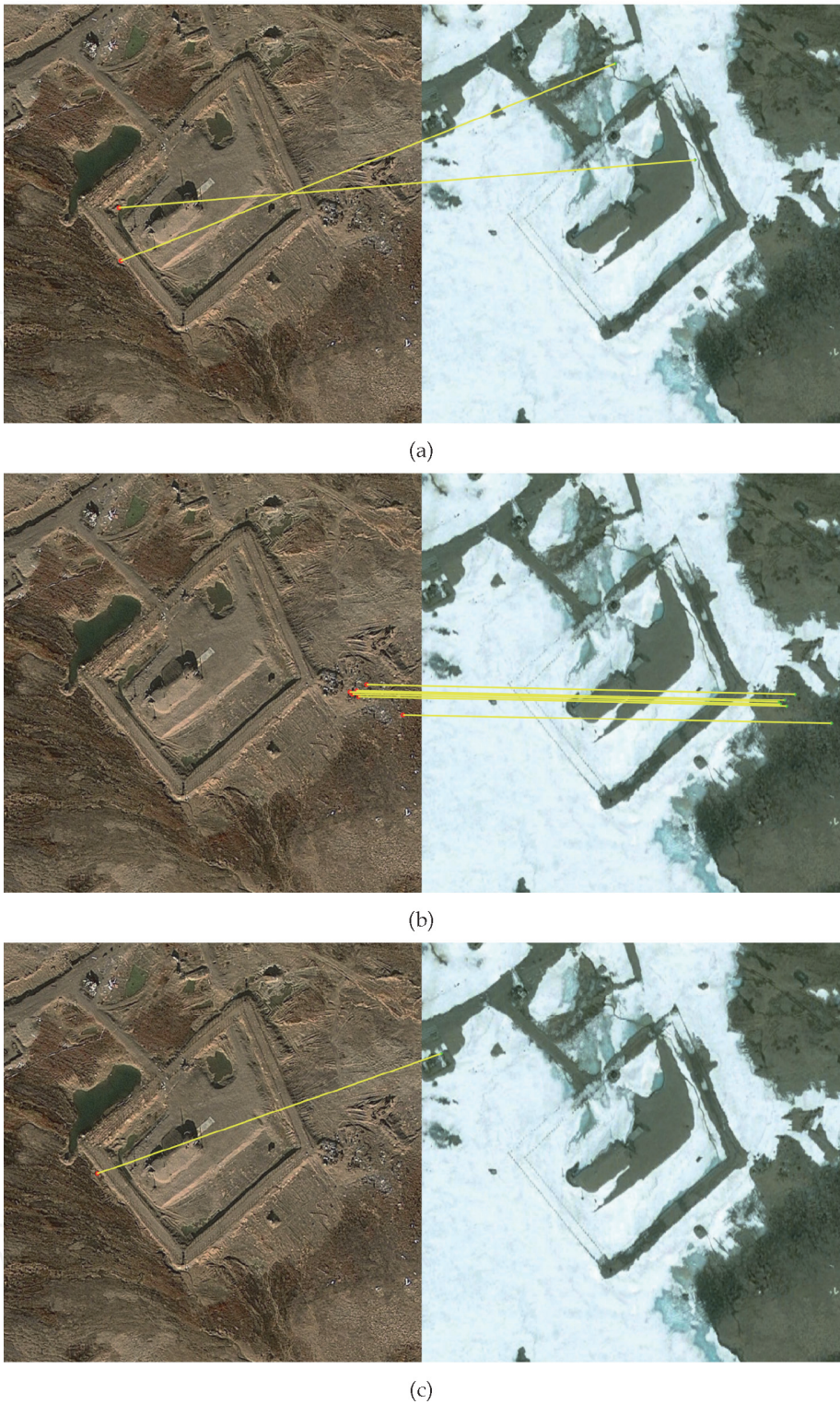


Figure 6.
Image pair (I_0, I_2) matching result for: (a) SURF, (b) KAZE, and (c) MSER method.

of the pair (I_0, I_2) were located. When comparing a relatively similar pair (I_2, I_3) , it appeared that all algorithms indicated mismatches or lack thereof, with KAZE and MSER indicating two correct matches each (Table 5).

In general, the KAZE method proved to be the most effective, while the ORB method showed the least processing efficiency of the set I. Due to the lack of any pair (I_0, I_2) , (I_0, I_3) and (I_2, I_3) matches, no graphical results are presented for the ORB method. A potential cause of a lack of pair (I_2, I_3) matches is a large contrast change, characteristic of the occurrence of acquisition interference, such as the fog visible in the image I_3 .

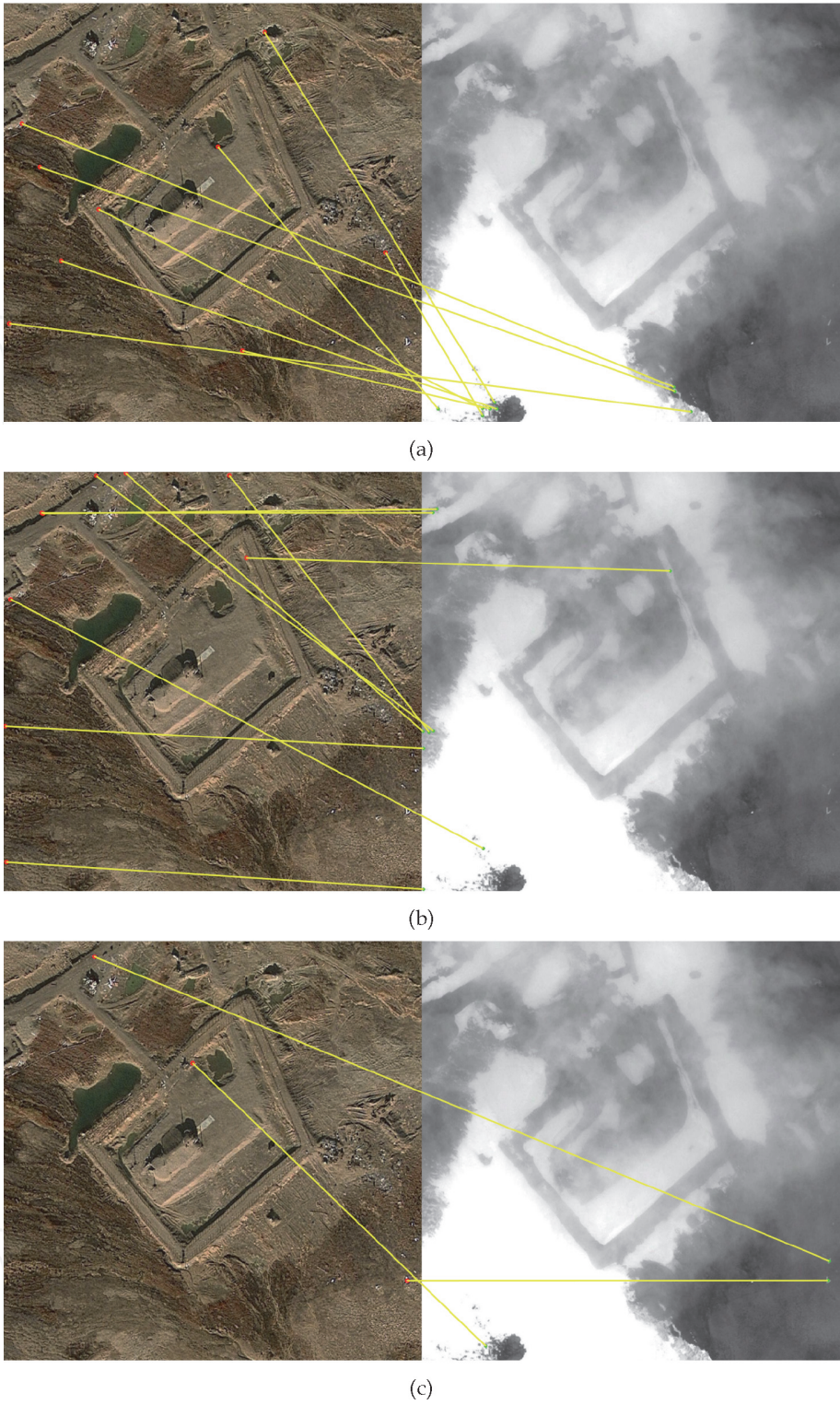


Figure 7.
Image pair (I_0, I_3) matching result for: (a) SURF, (b) KAZE, and (c) MSER method.

4.3 Effect of contrast change on the number of the detected features

The research focused on the analysis of the effect of contrast change on the number of features detected in the image. For this purpose, the contrast of the image I_0 had been gradually reduced until a uniform colour throughout the image was obtained. Afterwards, the transformed set was further analysed. SURF, MSER, ORB and KAZE methods were used again. **Figure 9** shows the cumulative results of this study.

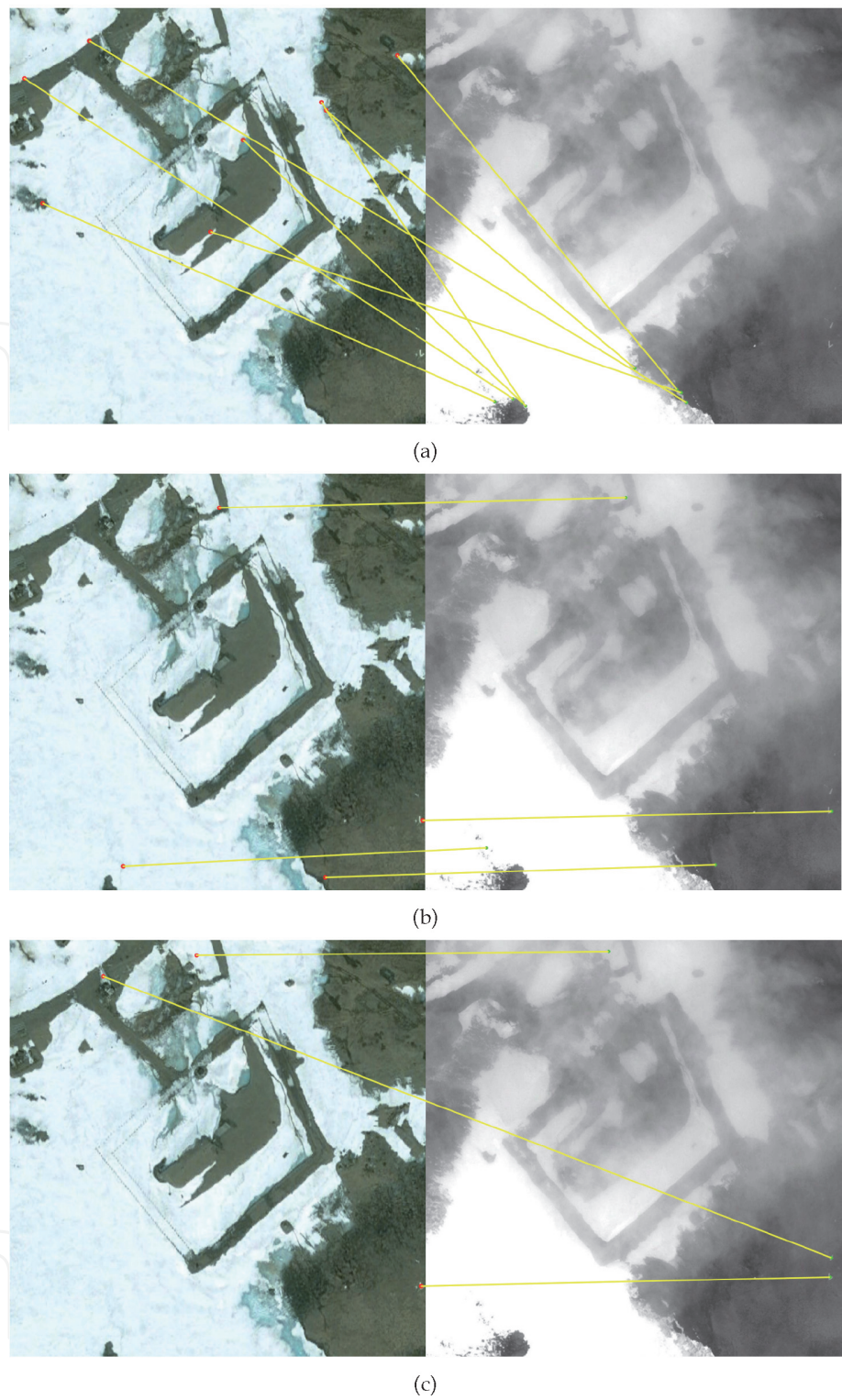


Figure 8.
Image pair (I_2, I_3) matching result for: (a) SURF, (b) KAZE, and (c) MSER method.

	SURF	KAZE	MSER	ORB
Correct matches	9	489	2	11
Mismatches	0	0	1	0
Percentage of correct matches	100%	100%	67%	100%

Table 2.
Image pair (I_0, I_1) matching results.

	SURF	KAZE	MSER	ORB
Correct matches	0	6	0	0
Mismatches	2	0	1	0
Percentage of correct matches	0%	100%	0%	0%

Table 3.
Image pair (I_0, I_2) matching results.

	SURF	KAZE	MSER	ORB
Correct matches	0	2	1	0
Mismatches	9	7	2	0
Percentage of correct matches	0%	29%	33%	0%

Table 4.
Image pair (I_0, I_3) matching results.

	SURF	KAZE	MSER	ORB
Correct matches	0	3	2	0
Mismatches	8	1	1	0
Percentage of correct matches	0%	75%	67%	0%

Table 5.
Image pair (I_2, I_3) matching results.

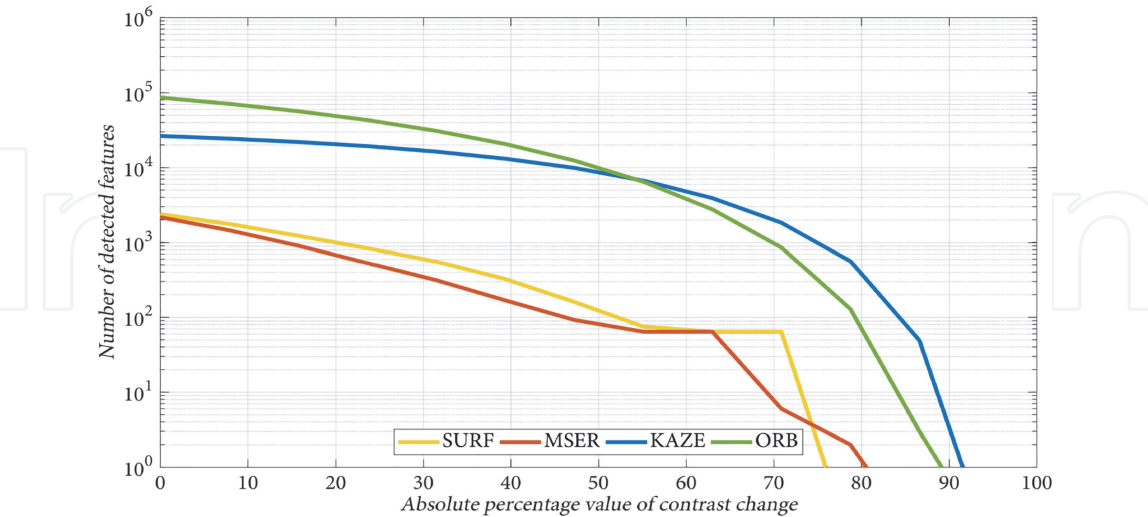


Figure 9.
Effect of contrast change on the number of the detected image features.

On the basis of the results obtained, it can be concluded that the number of features detected by the examined methods decreases with image contrast reducing, which results in a smaller statistical sample processed in each subsequent step of these methods. This may be the cause of the lower matching efficiency of the methods considered for images that are significantly different from each other.

5. Conclusions and final remarks

The results of the algorithms presented in the literature are usually related to images that are fragments of source images, i.e. have their qualitatively identical counterparts in Ref. images. In the analysed cases, the differences between the reference image stored in the memory of the navigation system and that recorded by the sensor are significant. As a result, there are certain consequences that often prevent the image representing the same field object from being effectively matched. This is due to real environmental conditions and restrictions on obtaining information. The measurement system parameters and the quality of the images taken have a direct impact on the number of detected features. For example, the lack of complete information about the accuracy of field object's image mapping makes it impossible to properly select the size of the filters. This results in the detection of objects that are completely irrelevant to the issue considered, such as bushes, leaves or grass blades, which are highly variable over time. Consequently, it has a significant impact on the performance of individual algorithms.

The study concluded that the use of statistical algorithms such as RANSAC improves the effectiveness of the selected methods. However, the results obtained strongly depend on the size of the set taken into consideration and the match/mismatch ratio. Therefore, in the terrain image processing, it is necessary to conduct an analysis of the informational characteristics of the examined objects and the conditions of acquisition. This allows for extracting characteristic points whose description does not significantly change due to atmospheric conditions.

The results of the simulation tests enable a general conclusion that the methods considered are often insufficient to determine the coordinates of a UAV/CM flying under unfavourable environmental conditions. The greatest development potential, in the context of the implementations examined in this work, is characterised by methods based on anisotropic diffusion, which in the course of simulation studies showed the highest effectiveness. Therefore, it seems justified to focus the research effort on further development of new image processing methods within the group of anisotropic diffusion methods. In particular, it is proposed to take the informative character of terrain images as determinants of the input parameters of the designed processing methods into account, to apply pre-processing methods aimed at decimation of the input data, their segmentation and determination of the main components, and to extend the definition of the designed methods with additional criteria increasing the effectiveness of detection and image feature matching. The newly developed methods should be aimed at the improvement of feature detection efficiency in terrain images and the selection of processing parameters taking into account environmental conditions as well as limitations and conditions in the measurement system.

Acknowledgements

This work is financed by the National Centre of Research and Development of the Republic of Poland as part of the scientific research program for the defence and security named *Future Technologies for Defence – Young Scientist Contest* (Grant No. DOB-2P/03/06/2018).

Conflict of interest

The authors declare no conflict of interest.

IntechOpen

IntechOpen

Author details

Piotr Turek, Stanisław Grzywiński and Witold Bużantowicz*
Military University of Technology, Warsaw, Poland

*Address all correspondence to: witold.buzantowicz@wat.edu.pl

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Boozer DD, Fellerhoff JR. Terrain-Aided Navigation Test Results in the AFTI/F-16 Aircraft. *Navigation – Journal of The Institute of Navigation*. 1988;35(2):161–175. DOI: 10.1002/j.2161-4296.1988.tb00949.x
- [2] Enns R. Terrain-aided navigation using the Viterbi algorithm. *Journal of Guidance, Control, and Dynamics*. 1995; 18(6):1444–1449. DOI: 10.2514/3.21566
- [3] Han Y, Wang B, Deng Z, Fu M. An improved TERCOM-based algorithm for gravity-aided navigation. *IEEE Sensors Journal*. 2016;16(8):2537–2544. DOI: 10.1109/JSEN.2016.2518686
- [4] Hua Z, Xiulin H. A height-measuring algorithm applied to TERCOM radar altimeter. In: *Proc. of the 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*; 20–22 August 2010; Chengdu (China). New York: IEEE, 2010. p. (V5–43)-(V5–46). DOI: 10.1109/ICACTE.2010.5579215
- [5] Wei E, Dong C, Liu J, Tang S. An improved TERCOM algorithm for gravity-aided inertial navigation system. *Journal of Geomatics*. 2017;42(6):29–31. DOI: 10.14188/j.2095-6045.2016190
- [6] Naimark L, Webb H, Wang T. Vision-Aided Navigation for Aerial Platforms. In: *Proc. of the ION 2017 Pacific PNT Meeting*; 1–4 Mai 2017; Honolulu (USA). Manassas: ION, 2017. p. 70–76. DOI: 10.33012/2017.15051
- [7] Yang C, Vadlamani A, Soloviev A, Veth M, Taylor C. Feature matching error analysis and modeling for consistent estimation in vision-aided navigation. *Navigation*. 2018;65:609–628. DOI: 10.1002/navi.265628
- [8] Carr JR, Sobek JS. Digital Scene Matching Area Correlator (DSMAC). In: *Proc. of 24th Annual Technical Symposium, SPIE 0238, Image Processing for Missile Guidance*; 23 December 1980; San Diego (USA). Bellingham: SPIE, 1980. DOI: 10.1117/12.959130
- [9] Irani GB, Christ JP. Image processing for Tomahawk scene matching. *Johns Hopkins APL Technical Digest*. 1994;15 (3):250–264.
- [10] Turek P, Bużantowicz W. Image matching constraints in unmanned aerial vehicle terrain-aided navigation. In: *Proc. of the 2nd Aviation and Space Congress*; 18–20 September 2019; Cedzyna (Poland). p. 206–208.
- [11] Brown LG. A survey of image registration techniques. *ACM Computing Surveys*. 1992;24(4):325–376. DOI: 10.1145/146370.146374
- [12] Zitová B, Flusser J. Image registration methods: A survey. *Image and Vision Computing*. 2003;21(11): 977–1000. DOI: 10.1016/S0262-8856 (03)00137-9
- [13] Bouchiha R, Besbes K. Automatic Remote-Sensing Image Registration Using SURF. *International Journal of Computer Theory and Engineering*. 2013;5(1):88–92. DOI: 10.7763/IJCTE.2013.V5.653
- [14] Kashif M, Deserno TM, Haak D, Jonas S. Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment. *Computers in Biology and Medicine*. 2016;68:67–75. DOI: 10.1016/j.compbimed.2015.11.006
- [15] Lindeberg T. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*. 1994;21(2):224–270. DOI: 10.1080/757582976
- [16] Löwe DG. Distinctive Image Features from Scale-Invariant

Keypoints. *International Journal of Computer Vision*. 2004;60:91–110. DOI: 10.1023/B:VISI.0000029664.99615.94

[17] Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*. 2008;110(3):346–359. DOI: 10.1016/j.cviu.2007.09.014

[18] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: *Proc. of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 8–14 December 2001; Kauai (USA). p. (I-511)-(I-518). DOI: 10.1109/CVPR.2001.990517

[19] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. In: *Proc. of the 13th International Conference on Computer Vision*; 6–13 November 2011; Barcelona (Spain). p. 2564–2571. DOI: 10.1109/ICCV.2011.6126544

[20] Rosten E, Drummond T. Machine Learning for High-Speed Corner Detection. In: *Leonardis A, Bischof H, Pinz A, editors. Proc. of the 9th European Conference on Computer Vision 2006 – Lecture Notes in Computer Science*, vol. 3951. Berlin-Heidelberg: Springer; 2006. p. 430–443. DOI: 10.1007/11744023_34

[21] McIlroy P, Rosten E, Taylor S, Drummond T. Deterministic sample consensus with multiple match hypotheses. In: *Proc. of the 21st British Machine Vision Conference*; 31 August – 3 September 2010; Aberystwyth (UK). pp. 111.1–111.11. DOI: 10.5244/C.24.111

[22] Calonder M, Lepetit V, Strecha C, Fua P. BRIEF: Binary Robust Independent Elementary Features. In: *Daniilidis K, Maragos P, Paragios N, editors. Proc. of the 11th European Conference on Computer Vision 2010 – Lecture Notes in Computer Science*, vol. 6314. Berlin-Heidelberg: Springer; 2010.

p. 778–792. DOI: 10.1007/978-3-642-15561-1_56

[23] Alcantarilla PF, Bartoli A, Davison AJ. KAZE Features. In: *Fitzgibbon A, Lazebnik S, Perona P, Sato Y, Schmid C, editors. Proc. of the 13th European Conference on Computer Vision 2010 – Lecture Notes in Computer Science*, vol. 7577. Berlin-Heidelberg: Springer; 2012. p. 214–227. DOI: 10.1007/978-3-642-33783-3_16

[24] Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1990;12(7):629–639. DOI: 10.1109/34.56205

[25] Weickert J. Efficient image segmentation using partial differential equations and morphology. *Pattern Recognition*. 2001;34:1813–1824. DOI: 10.1016/S0031-3203(00)00109-6

[26] Charbonnier P, Blanc-Feraud L, Aubert G, Barlaud M. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*. 1997;6(2): 298–311. DOI: 10.1109/83.551699

[27] Matas J, Chum O, Urban M, Pajdla T. Robust wide baseline stereo from maximally stable extremal regions. In: *Proc. of the 13th British Machine Vision Conference*; 2–5 September 2002; Cardiff (UK). pp. 384–396.

[28] Chaumette F. Image moments: a general and useful set of features for visual servoing. *IEEE Transactions on Robotics*. 2004;20(4):713–723. DOI: 10.1109/TRO.2004.829463

[29] Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*. 1981;24(6):381–395.

[30] Wang Z, Bovik AC. Mean squared error: love it or leave it?. IEEE Signal Processing Magazine. 2009;26(1):98–117. DOI: 10.1109/MSP.2008.930649

[31] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing. 2004;13(4):600–612.

[32] Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM. In: Proc. of the 20th IAPR International Conference on Pattern Recognition; 23–26 August 2010; Istanbul (Turkey). p. 2366–2369. DOI: 10.1109/ICPR.2010.579